

Mixed Models for Short-Run Forecasting of Electricity Prices: Application for the Spanish Market

Carolina García-Martos, Julio Rodríguez, and María Jesús Sánchez

Abstract—Short-run forecasting of electricity prices has become necessary for power generation unit schedule, since it is the basis of every profit maximization strategy. In this article a new and very easy method to compute accurate forecasts for electricity prices using mixed models is proposed. The main idea is to develop an efficient tool for one-step-ahead forecasting in the future, combining several prediction methods for which forecasting performance has been checked and compared for a span of several years. Also as a novelty, the 24 hourly time series has been modelled separately, instead of the complete time series of the prices. This allows one to take advantage of the homogeneity of these 24 time series. The purpose of this paper is to select the model that leads to smaller prediction errors and to obtain the appropriate length of time to use for forecasting. These results have been obtained by means of a computational experiment. A mixed model which combines the advantages of the two new models discussed is proposed. Some numerical results for the Spanish market are shown, but this new methodology can be applied to other electricity markets as well.

Index Terms—Design of experiments, electricity markets, forecasting, marginal price, time series analysis.

I. INTRODUCTION

ELECTRICITY is nowadays traded under competitive rules, like other *commodities*. But electricity presents some characteristics that make it different, since it cannot be stored and any non-served demand is lost. These special features are responsible for the extremely volatile and largely unpredictable behaviour of its price.

In competitive markets, there are several ways to trade electricity: bilateral contracts, the *pool* and forward markets and options, which are well developed in some electricity markets like the EEX in Germany. As far as bilateral contracts are concerned, an interesting question is how to reduce the risks that they imply. This can be done by forecasting electricity prices with a horizon that covers, at least, the length of the contract, usually one year, which means long-term forecasting. In the *pool*, both the generating companies and the consumers submit to the market operator their respective generation and consumption bids for each hour of the forthcoming day. In the Spanish market, once the market operator has sorted out the bidding prices for generation

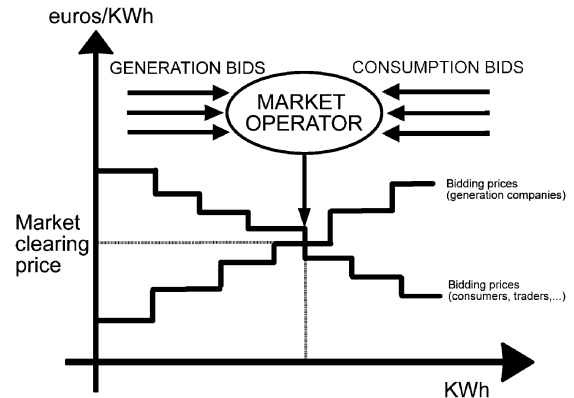


Fig. 1. Market clearing price, Spanish market.

or consumption bids respectively, the marginal price is defined as the bid submitted by the last generation unit needed to satisfy the whole demand. This mechanism to obtain what is also known as the market clearing price is shown in Fig. 1.

Having short-term accurate forecasts helps the producers schedule the production of their units for profit maximization. Certain risk is assumed by the producer but, the more accurate the forecasts are, the lower the risk. A generating company can better decide its bidding price when having accurate one-day-ahead forecasts. A powerful tool for short-run forecasting is the basis on which every bidding-rule stands ([1], [2]). Bearing this in mind, the availability of adequate models to predict next-day electricity prices is of great interest.

Just a few years ago, demand was the only factor predicted in centralized markets. Nevertheless, forecasting electricity prices is relatively recent, but several techniques have been applied and can be divided into two main groups: Neural networks and time series models. Neural networks have been used by Ramsay *et al.* [3], Szkuta *et al.* [4] and Rodríguez *et al.* [5]. Nicholaisen *et al.* [6] have produced forecasts for electricity prices using neural networks filtering the non-linearity of the prices with Fourier and Hartley transforms. Models based on time series have also been applied to electricity prices. Nogales *et al.* [7] applied transfer functions and dynamic regression to forecast electricity prices. Contreras *et al.* [8] forecasted electricity prices of the Spanish and Californian markets by applying ARIMA models. Troncoso *et al.* [9] compared the kWNN (k Weighted Nearest Neighbours) technique with dynamic regression. Crespo-Cuaresma *et al.* [10] have suggested a group of univariate models to predict electricity prices in the Leipzig market, the most important spot market in Germany. Conejo *et al.* [11] compared several methods including wavelet approximation, ARIMA models and neural networks. Nogales *et al.*

C. García-Martos and M. J. Sánchez are with E.T.S. Ingenieros Industriales, Universidad Politécnica de Madrid, C/ José Gutiérrez Abascal 2, Madrid 28006, Spain (e-mail: cgarciam@etsii.upm.es).

J. Rodríguez is with Facultad de Ciencias Económicas, Universidad Autónoma de Madrid, Madrid 28007, Spain.

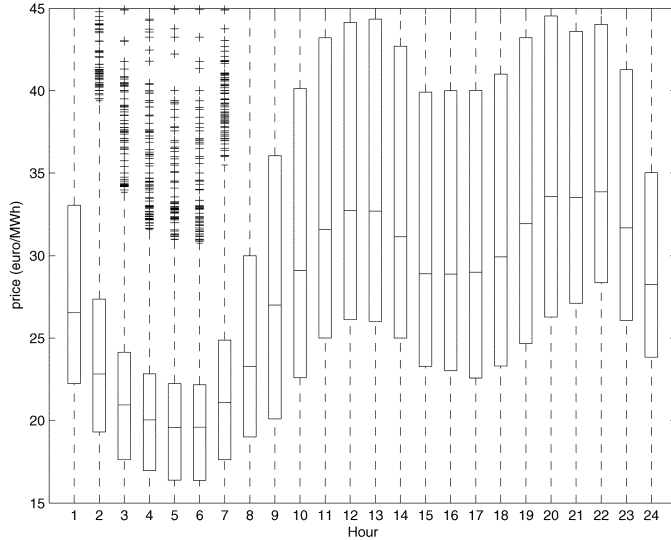


Fig. 2. Boxplot for hourly prices.

[12] forecasted the prices in the PJM interconnection through transfer functions, showing that the inclusion of explanatory variables does not significantly reduce the prediction errors.

In this article we propose a simple but accurate method to compute one-day-ahead forecasts of electricity prices. We propose two new models, both of which deal with the 24 hourly time series of electricity prices instead of the complete one. We also try to determine the appropriate length of the time series used to build the forecasting ARIMA models. We carry out a computational experiment to determine the combination of Model and Length with the best “global performance.”

This paper is organized as follows. In Section II the methodology and the computational implementation are described. In Section III a study of forecasting errors is carried out from two different points of view: descriptively and applying *loess* (locally weighted regression). In Section IV we develop the design of experiments. Section V presents numerical results for the Spanish market. In Section VI some conclusions are provided.

II. METHODOLOGY AND COMPUTATIONAL IMPLEMENTATION

The methodology will be described taking the Spanish market as an example for illustration.

The main idea is to compute accurate forecasts building a mixed model that combines the advantages of the ones used to build it.

In this paper two new models to deal with this problem are proposed. Both of them deal with the 24 hourly time series of electricity prices.

A brief descriptive analysis of the prices corresponding to the period 1998–2003 has been made. Fig. 2 presents a boxplot of the hourly prices for the period under study.

A boxplot is a graphical representation of a distribution, built to show its most important features and the outliers. The limits of the rectangle are quantiles 25 and 75 (Q_{25} and Q_{75}). The position of the median, Q_{50} , is indicated by drawing a line. By construction, 50% of the data in the sample is inside the rectangle, and 75% of the data is smaller than Q_{75} . Besides, Q_{50} is

the median, so it is the center of the sample. Admissible values for the upper and lower limit (UL and LL respectively) are calculated. These limits can be used to identify outliers.

Boxplots are especially useful for providing a general idea about the distribution of a variable.

In this work, we have used boxplots to study the distribution of hourly prices (level and variability), and also for having a better picture of forecasting errors. Not only the level of the prices but also their variability depends on the corresponding hour of the day. This conclusion can be extended to other markets since it is a consequence of the instantaneous effect of demand on price.

The prediction errors obtained by separately studying and modelling each of these 24 hourly time series, far from being affected negatively by the loss of information, are reduced. The larger homogeneity of the 24 hourly time series in comparison with the complete one, as well as the fact that 24 one-step-ahead forecasts are calculated everyday instead of 24 individual forecasts with prediction horizons varying from 1 to 24, will allow an improvement of the accuracy of the forecasts.

Two new different models are proposed. The first one, which we will refer to, from now on as Model 24, forecasts electricity prices for each of the 24 hours of the next day using the ARIMA models built for each of the 24 time series. So, to produce a forecast with Model 24 for tomorrow’s H^{th} hour, we use the model estimated for this hour with the previous L complete weeks (7-day week, considering both workday and weekend data).

The second model computes the forecasts for working days using the 24 workday time series and the forecasts of the prices in weekends with the weekend data. This second model is hereafter referred to as Model 48, because of the number of series with which it works (24 series for working days and 24 for weekends). To produce a forecast for tomorrow’s H^{th} hour (if it is a weekday), we use the estimation of the model for this hour built with the previous L weeks (considering 5-day weeks, only weekday data). On the other hand, to produce a forecast for tomorrow’s H^{th} hour (if it is Saturday or Sunday), we use the estimation of the model for this hour built with the previous L weeks (considering a 2-day week, only the weekend data).

Besides, the disaggregation of the price time series into 24 hourly time series allows the use of ARIMA methodology in better conditions, because the frequency of the data is reduced.

A computational experiment has been carried out to determine which model leads to more accurate forecasts. Some relevant factors in short term prediction have been included. We have thus obtained the appropriate length of the time series used to build the forecasting models. Until now a default length of two or three months had been used to build the models, but no published study analyzes this issue.

It is of interest to analyze the sources of variability affecting the prediction error. We thus chose the following factors and levels.

- Model: Model 48 or Model 24.
- Length of the time series used to build the forecasting model:
 - 8, 12, 16, 20, 24, 28, 32, 44, 52, and 80 weeks.

Forecasts have been computed for all the prices in the period from 1 January 1998 to 31 December 2003, both for Model 24

and Model 48, for the ten possible levels of the factor “length of the time series used to build the model.” The size of the sample as well as its representativity, as it includes the years 1998–1999 in which the variability of the prices is smaller because of the initialization of the competitive markets, years 2000–2003 when larger variability in the prices can be appreciated and the great increase in the prices that occurred at the end of 2001 and the beginning of 2002, will allow inferences and extend the conclusions reached in forecasting prices in the future. We want the combination of Model and Length that best fits in “global terms” for a very long period of time in which different patterns in level and variability of the prices were registered. This mixed model would be able to produce accurate forecasts for future periods, depending neither on the level of the prices nor on their variability.

We would like to highlight the great number of models to be identified and estimated (bear in mind that the models are refitted everyday). Forecasting prices for one day using a different multiplicative ARIMA model for each hour (as we propose) implies the identification and estimation of 24 models. If the objective is extended to computing forecasts for the prices in one week, $24 \times 7 = 168$ models must be identified and estimated. Given our ambitious objectives, trying to compute forecasts for the six years considered ($6 \times 365 \times 24 = 52854$ hours) with the $20 = 2 \times 10$ possible combinations of the levels of the two factors considered (two levels for the factor Model and ten levels for the factor Length of the time series) required more than 1051200 ($= 6 \times 365 \times 24 \times 20$) models to be identified and estimated.

With the huge amount of models to identify, more than 1000000, it is impossible to manual fit of all of them using the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF), and requires the automatization of the procedure.

Software like SCA or EViews can deal with automatic identification of ARIMA models and also includes the option of outlier identification and intervention, but none of them are free and expensive licenses are required to use them. In this paper we have used TRAMO (“Time series Regression with Arima noise, Missing observations and Outliers”) [13], a software developed by G. Caporello and A. Maravall (of the Bank of Spain) for the estimation and subsequent forecasting with ARIMA models. Identification and intervention of outliers, as well as estimation of models can be done with an automatic procedure. Models are selected using the Bayesian Information Criteria (BIC). This criterion takes into account both the likelihood of the model (by means of its residual variance) and the parsimony of the model (including a term that penalizes models with a great number of parameters).

The expression of the Bayesian Information Criteria is

$$BIC = n \log(\hat{s}_R^2) + k \log(n)$$

where n is the length of the time series used to estimate the model, \hat{s}_R^2 is the residual variance, and k is the number of parameters estimated.

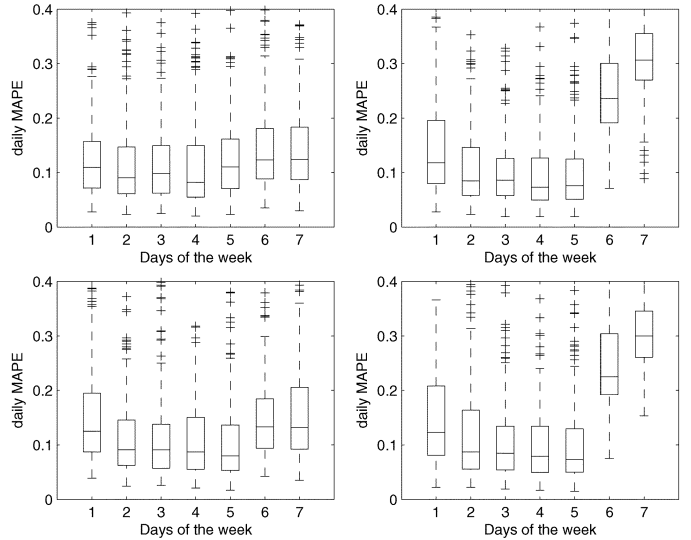


Fig. 3. Boxplot forecasting errors. Models 24, 48 (first and second column respectively). Lengths 44, 80 weeks (first and second row, respectively).

Capabilities of TRAMO are similar to those in SCA or EViews, but the main advantage of TRAMO is that it is free software, it can be downloaded from the Web (<http://www.bde.es/servicio/software/softwaree.htm>) and it is easy to use for applied statisticians, engineers, etc.

Finally, it is important to bear in mind that this paper is not restricted to obtaining forecasts for several days or weeks with small prediction errors. We have computed forecasts for all the hours in a long period of time (six years) and it is of our interest to find out the appropriate levels of the two factors considered (Model and Length of the series) in order to minimize prediction errors in “global terms.”

III. ANALYSIS OF FORECASTING ERRORS

Once the forecasts had been computed for all possible combinations of factor levels, it is of great interest to have a prior idea of the convenience of using Model 24 or Model 48 to forecast prices and whether this decision depends on the length of the time series used to build the model or on a working day versus weekend.

Some conclusions can be drawn from a brief descriptive analysis of the prediction errors: The prediction errors decrease as the length of the time series increases, and this occurs up to a length of around 44 weeks. Building the models with a longer series produces less accurate forecasts for weekends and not significantly smaller prediction errors for weekdays. It can also be observed that Model 24 provides a better fitting for weekends while Model 48 does likewise for working days. This can be observed in Fig. 3, which shows the boxplot of the prediction errors for Models 24 and 48 and lengths 44 and 80 weeks for all the days in the week.

This brief descriptive analysis of the prediction errors indicates that it would be reasonable to carry out the design of experiments separately for working days and weekends. Besides, a rough a priori estimate of the model and the length of the time series for both weekends and working days can be obtained from

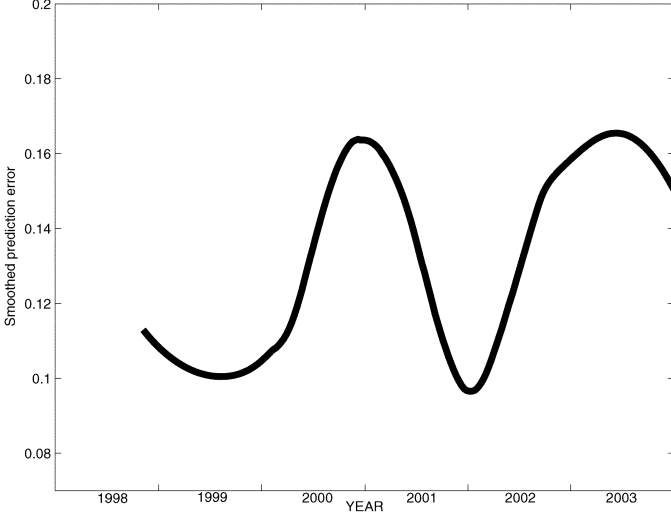


Fig. 4. Smoothed prediction error for the whole period considered. Model 24 for weekends and Model 48 for weekdays. Length equal to 44 weeks in both cases. First forecast computed for the 45th week in 1998 since the first 44 weeks are used to identify and estimate the model.

Fig. 3. Thus, a nonparametric estimation of the conditional mean of the errors that takes into account the time evolution is shown in Fig. 4. This estimation has been computed by means of *loess* (Locally weighted regression), [14], and using Model 24 for weekends and Model 48 for weekdays, and Length equal to 44 weeks in both cases.

Some preliminary conclusions can be drawn from this analysis. Forecasts computed using Model 48 are more accurate for working days. Nevertheless, for one-day-ahead forecasts of electricity prices on weekends, Model 24 leads to smaller prediction errors. The optimal length to build the models seems to be around 44 weeks.

These preliminary conclusions can now be used as follows. Since the two new models, 24 and 48, fit better for weekends and working days respectively, in the next section the design of experiments will be carried out separately for working days and weekends.

The average daily prediction error for the full period considered (November 1998-December 2003) is 12.6%. Bearing in mind that forecasts have been calculated for a representative and large period of time, the results, in terms of prediction error, reflect the accuracy of the new mixed model proposed.

IV. DESIGN OF EXPERIMENTS

The main objective of this paper is to develop a simple but efficient tool to compute one-day-ahead forecasts for electricity prices.

To do this we will determine whether it is useful or not to identify, estimate and forecast with different models for working days and weekends, as well as the optimal number of observations in the time series used to build the models. To select the appropriate Model and Length we carried out a computational experiment. The day for which the forecasts are being computed for is included as a block, which means not taking into account the interactions between the block and the other

two factors (Model and Length). By doing so, the effect of the day is eliminated, since it is not of interest to us. This is because if the price in one day is unexpected, the forecast will not be accurate whatever the combination of the levels of the factors considered. Also, the possible correlation between forecasting errors is eliminated when including the day as a block. The nonexistence of correlations in the variable under study is one of the hypotheses assumed for the design of experiments.

The variable under study is the logarithm of the average daily prediction error. The factors considered are the model (24 or 48) and the length of the time series (8, 12, 16, 20, 24, 28, 32, 44, 52, and 80 weeks). As mentioned above the day is included as a block. The equation for the linear model of the computational experiment is

$$y_{ijt} = \mu + \alpha_i + \beta_j + \gamma_t + (\alpha\beta)_{ij} + u_{ijt}, \quad u_{ijt} \sim N(0, \sigma)$$

where μ is a global effect, i.e., the average level of the response (logarithm of the prediction error).

α_i is the main effect of the model. It measures the increase/decrease of the average response for model i with respect to the average level.

β_j is the main effect of the length of the time series. It measures the increase/decrease of average response for length j with respect to the average level.

γ_t is likewise the main effect of the block (the day in this particular case).

The term $(\alpha\beta)_{ij}$ measures the difference between the expected value of the response and the one computed using a model that does not include the interactions.

The random effect, u_{ijt} , includes the effect of all other causes.

The response, y_{ijt} , which is the logarithm of the Daily MAPE, has been calculated as follows:

$$y_{ijt} = \log(\text{DailyMAPE}) = \log \left(\frac{1}{24} \sum_{h=1}^{24} \frac{|\widehat{p}_t^h - p_t^h|}{p_t^h} \right)$$

where the forecast of the price \widehat{p}_t^h for day t in hour h has been calculated using model i and estimating the ARIMA model using the previous j observations. In the well known formulation of the $ARIMA(p, d, q) \times (P, D, Q)$ models

$$\phi_p(B)\Phi_P(B^s)\nabla^d\nabla_S^D p_t^h = \theta_q(B)\Theta_Q(B^s)a_t \quad (1)$$

where $\phi_p(B) = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)$, $\Phi_P(B^s) = (1 - \Phi_1 B^s - \Phi_2 B^{2s} - \dots - \Phi_P B^{Ps})$, $\nabla^d = (1 - B)^d$, $\nabla_S^D = (1 - B^s)^D$, $\theta_q(B) = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q)$ and $\Theta_Q(B^s) = (1 - \Theta_1 B^s - \Theta_2 B^{2s} - \dots - \Theta_P B^{Ps})$ $B p_t^h = p_{t-1}^h$.

The roots of $\phi_p(B)$, $\Phi_P(B^s)$, $\theta_q(B)$ and $\Theta_Q(B^s)$ must be outside the unit circle for stationarity.

The forecasts, \widehat{p}_t^h , have been calculated minimizing the expression

$$E \left[(p_t^h - \widehat{p}_t^h)^2 \right]$$

TABLE I
ANOVA FOR WEEKENDS

Source	Sum sq	D.f.	mean sq	F-stat	p-Val
A: Day	190.89	82	2.32	29.24	0.0
B: Length	0.79	9	0.08	0.98	0.45
C: Model	142.17	1	142.17	1785.91	0.0
Residual	121.56	1527	0.08		
TOTAL	452.88	1619			

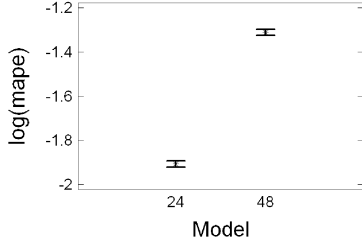


Fig. 5. Main effect model (weekends). Means and 95% Bonferroni intervals.

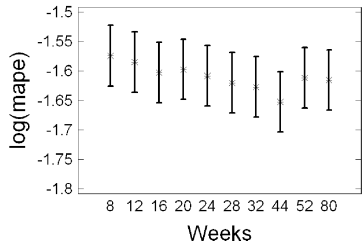


Fig. 6. Main effect length (weekends). Means and 95% Bonferroni intervals.

TABLE II
ANOVA FOR WORKING DAYS

Source	Sum sq	D. f.	Variance	F-stat	p-val
A: day	4310.46	506	8.52	194.3	0.0
B: length	12.21	9	1.35	30.95	0.0
C: model	16.69	1	16.69	380.80	0.0
Residual	414.88	9463	0.03		
TOTAL	4753.01	9979			

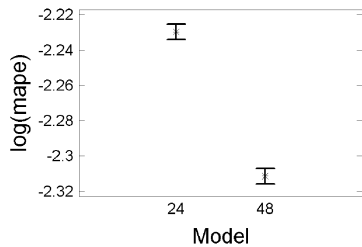


Fig. 7. Main effect model (weekdays). Means and 95% Bonferroni intervals.

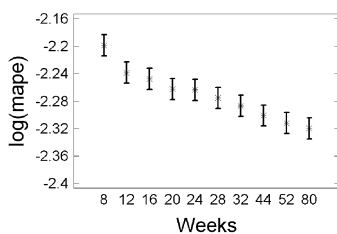


Fig. 8. Main effect length (weekdays). Means and 95% Bonferroni intervals.

which means that the best predictor is the expectancy of the conditional distribution

$$\hat{p}_t^h = E[p_t^h | p_{t-1}^h, p_{t-2}^h, \dots, p_1^h] = E[p_t^h | \mathbf{p}_{t-1}^h]$$

where $\mathbf{p}_{t-1}^h = (p_{t-1}^h, p_{t-2}^h, \dots, p_1^h)$.

So the expression for the response is

$$y_{ijt} = \log \left(\frac{1}{24} \sum_{h=1}^{24} \frac{|E[p_t^h | \mathbf{p}_{t-1}^h] - p_t^h|}{p_t^h} \right).$$

Bearing in mind the results of the previous section, working days and weekends are studied separately.

Table I shows the ANOVA table (Analysis of Variance, [15]) for the design of experiments of weekends. Interactions are not significant, but the main effects are.

An ANOVA table is based on variability decomposition. It includes the sum of squares of each effect as well as its variance (\hat{s}_{effect}^2), calculated dividing the sum of squares by the degrees of freedom. For each effect the F-statistic is calculated as follows:

$$F\text{-statistic} = \frac{\hat{s}_{effect}^2}{\hat{s}_{residual}^2}.$$

The null hypothesis $H_0 : E[\hat{s}_{effect}^2] = \sigma^2 = E[\hat{s}_{residual}^2]$ is rejected when the value of the F-statistic is significantly large.

Bonferroni adjustment has been applied to solve multiple comparisons problem, [15].

Figs. 5 and 6 show the main effects of Model and Length for weekends. The prediction error is significantly smaller using model 24, i. e. building the models using the electricity prices of the complete week. There is no significant difference between the prediction errors obtained building the models with prices of the previous 32, 44, or 80 weeks, as the intervals are overlapped; nevertheless, a lower level can be observed for 44 weeks. From this length on, daily prediction errors increase.

For weekdays the results of the Analysis of Variance (ANOVA) are shown in Table II.

Fig. 7 shows the convenience of Model 48 for working days. It can be observed in Fig. 8 that the daily prediction error decreases when we increase the number of weeks used to build the ARIMA models for the hourly time series. Using more than 44 weeks reduces the prediction error but not significantly (the intervals corresponding to 52 and 80 weeks are almost completely overlapped and the increase in the number of weeks is very large). Bearing this in mind and for simplicity, it would also be adequate to use the prior 44 weeks (as recommended for weekends) to build the models for weekdays. Although a decrease is shown in terms of prediction error when using the previous 80 weeks, it is not significant.

Until now and by default, the ARIMA models proposed to forecast short-term electricity prices used the previous 10–12 weeks, but no study about the effect of the length has been published.

The main difference between the methodology that we propose and previous ones is that we propose fitting different

models for each hour. The fact that prediction errors became always smaller when adding more information indicates that the 24 processes that generate hourly prices are much more homogeneous than the process that generates the complete time series. This is relevant in workdays, when there is always a decrease in terms of prediction error when computing forecasts with models that have been estimated with longer time series.

The main conclusions of the design of experiments are as follows.

- Convenience of using about 44 weeks prior to the day for which we forecast for weekends.
- Convenience of using 44 or more weeks for working days (although the mean is lower for 80 weeks there are not significant differences between them).
- For weekdays, better estimations of the parameters are obtained when including more data, this means that the generating processes of the 24-hourly time series are very homogeneous, much more than the process generating the complete time series.
- The development of a mixed model that forecasts for working days with model 48 and for weekends using model 24.

V. NUMERICAL RESULTS

The main goal of this paper is to develop an alternative simple model for obtaining accurate one-day-ahead electricity forecasts for electricity prices.

Using the new mixed model developed taking into account the conclusions of the design of experiments of Section IV, some results are now shown to evaluate the accuracy of the forecasts.

We want to remark that the great difference between our and previous works is that we have computed forecasts for every hour in the period of 1998–2003, and we have done this for the 20 possible combinations of the two factors under study and then selected the appropriate levels for these factors. That is why we could say that the mixed model that was developed is the one with the best “general performance” for the whole period considered.

In this section, we first show numerical results for eight specific weeks, six of them have been chosen to be comparable with results from previous works, ([8], [16]).

Our mixed model computes accurate forecasts in these specific weeks, but we want to highlight the global results we obtain for such a huge period of time (1998–2003).

The first week selected is the last one in May 2000 (25–31). Fig. 9 shows the real values and the forecasts. Daily mean errors for this week appear in Table III, as well as the prediction errors obtained with the model proposed by Contreras *et al.* [8]. The alternative mixed model proposed in this paper increases the accuracy of the forecasts, since for five days (out of seven) our daily mean error is smaller.

For illustration, Tables VIII and IX (in the Appendix) present the results of the estimation of the models used to compute the 24 hourly forecasts for May 26, 2000.

The second week for which results are shown is the last one in August 2000 (25–31). It is usually a low demand week.

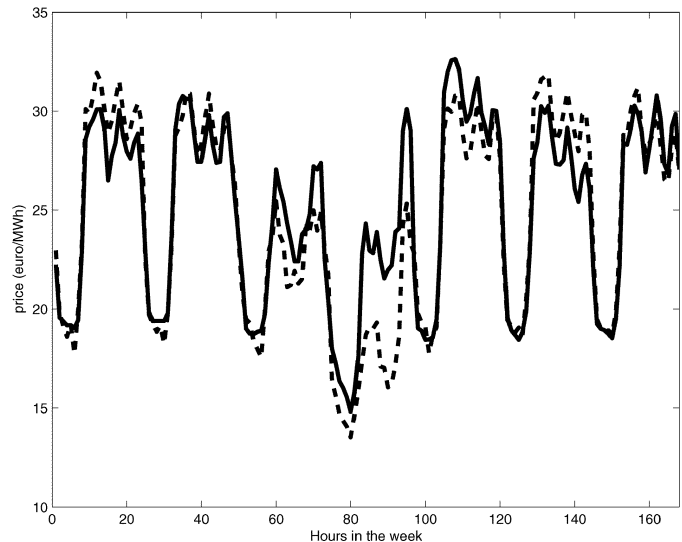


Fig. 9. Forecasts and real prices (25–31 May 2000).

TABLE III
PREDICTION ERRORS MAY 25–31 2000

Pred error	Mixed model proposed	Contreras et al [8]
Day 1	4.5 %	4.73 %
Day 2	1.9 %	4.13 %
Day 3	4.2 %	3.71 %
Day 4	16.5 %	6.84 %
Day 5	4.0 %	6.09 %
Day 6	5.6 %	6.09 %
Day 7	1.7 %	3.41 %

TABLE IV
PREDICTION ERRORS AUGUST 25–31 2000

Pred. error	Mixed model proposed	Contreras et al [8]
Day 1	4.8 %	4.3 %
Day 2	7.3 %	7.99 %
Day 3	5.4 %	4.57 %
Day 4	4.6 %	10.81 %
Day 5	5.1 %	6.12 %
Day 6	14.9 %	17.34 %
Day 7	7.2 %	6.05 %

TABLE V
MWE FOR SOME WEEKS IN 2002

WEEK	Mixed Model	ARIMA Model [16]
18 th -24 th Feb 2002	6.15 %	6.32 %
20 th -26 th May 2002	4.46 %	6.36 %
19 th -25 th Aug 2002	14.90 %	13.39 %
18 th -24 th Nov 2002	11.68 %	13.78 %

TABLE VI
PREDICTION ERRORS DECEMBER 15–21 2001

	Prediction error
Day 1	17.9 %
Day 2	8.7 %
Day 3	9.4 %
Day 4	5.7 %
Day 5	9.5 %
Day 6	5.3 %
Day 7	11.5 %

Fig. 10 shows the real prices and the forecasts. Table IV compares the prediction errors obtained applying our alternative methodology and the ones that appear in Contreras *et al.* [8].

TABLE VII
PREDICTION ERRORS APRIL 15–21 2002

Prediction error	
Day 1	2.2 %
Day 2	2.8 %
Day 3	2.5 %
Day 4	2.1 %
Day 5	2.5 %
Day 6	7.5 %
Day 7	7.6 %

TABLE VIII
GLOBAL RESULTS, MIXED MODEL. QUANTILES OF DAILY MAPE (1998–2003)

	Q ₂₅	Q ₅₀	Q ₇₅
Monday	7.9 %	11.8 %	19.4 %
Tuesday	5.7 %	8.5 %	14.4 %
Wednesday	5.7 %	8.6 %	12.5 %
Thursday	4.9 %	7.3 %	12.4 %
Friday	5.1 %	7.5 %	12.1 %
Saturday	8.7 %	11.8 %	17.7 %
Sunday	8.7 %	12.3 %	18.0 %

TABLE IX
MODELS AND DIAGNOSIS CHECKING

Hour	Model	Q-val	$h - n$	χ^2_{h-n}
Hour 5	$ARIMA(0, 1, 1) \times (0, 1, 1)_5$	44.31	34	48.6
Hour 10	$ARIMA(0, 1, 2) \times (0, 0, 0)_5$	25.06	35	49.8
Hour 14	$ARIMA(1, 0, 1) \times (0, 0, 0)_5$	43.24	34	48.6
Hour 18	$ARIMA(0, 1, 1) \times (0, 0, 0)_5$	37.50	35	49.8
Hour 22	$ARIMA(1, 1, 1) \times (0, 0, 0)_5$	38.52	34	48.6

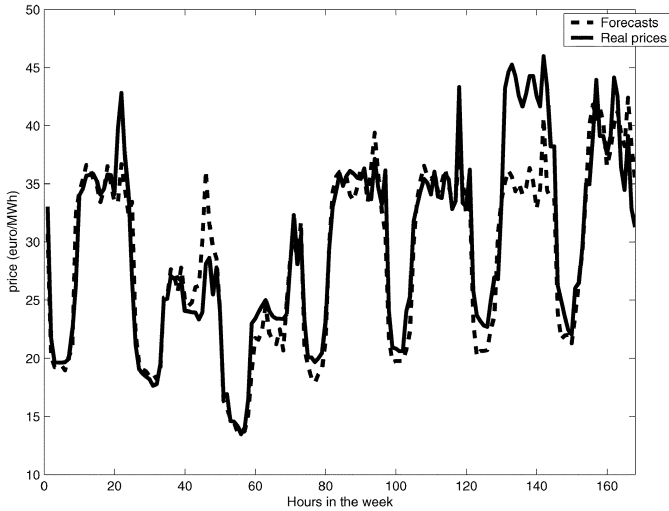


Fig. 10. Forecasts and real prices (25–31 August 2000).

Other four weeks selected are 18–24 February 2002, 20–26 May 2002, 19–25 August 2002, and 18–24 November 2002. We provide the Mean Week Error (MWE) obtained with our proposed mixed model and those obtained by Conejo *et al.* [16] with ARIMA methodology. Mean Week Error has been calculated as proposed in [16]

$$MWE = \frac{1}{168} \sum_{h=1}^{168} \frac{|p^h - \hat{p}^h|}{\bar{p}^h}.$$

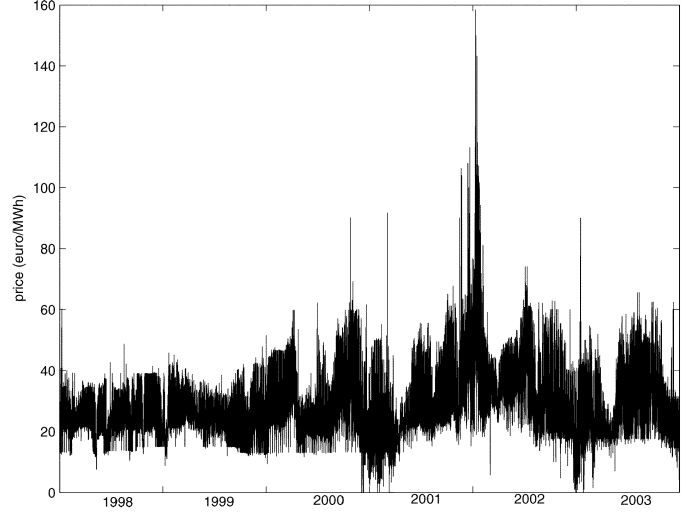


Fig. 11. Hourly prices (1 January 1998 – 1 December 2003).

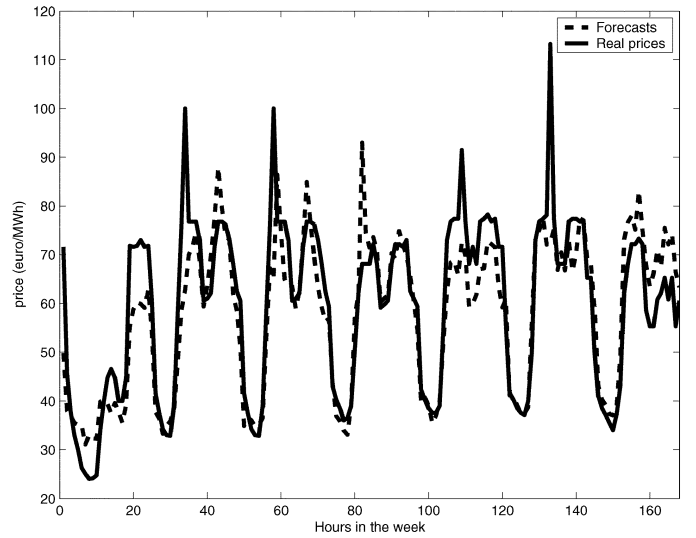


Fig. 12. Forecasts and real prices (15–21 December 2001).

The seventh week selected is the one before the last in December 2001 (15–21). It has been chosen because of the existence of a peak in demand, which influences the prices, as can be observed in Fig. 11, where we show the hourly prices in the whole period under study. The behaviour of the time series at the end of 2001 and the beginning of 2002 is clearly affected by the great increase in demand, caused by extremely low temperatures.

We would like to emphasize the accuracy of the forecasts of the mixed model. Even when the behaviour of the time series of the prices is rather unexpected, the alternative model proposed in this paper works properly, i.e., the prediction errors are small. The average prediction error for the whole week is 9.7%. The prediction errors appear in Table VI.

Fig. 12 shows the real prices and the forecasts for this week (15–21 December 2001).

Finally, the last week selected is the third one in April 2002. Since the ARIMA forecasts for one day are built using the 20 previous weeks, it is clear that the observations corresponding

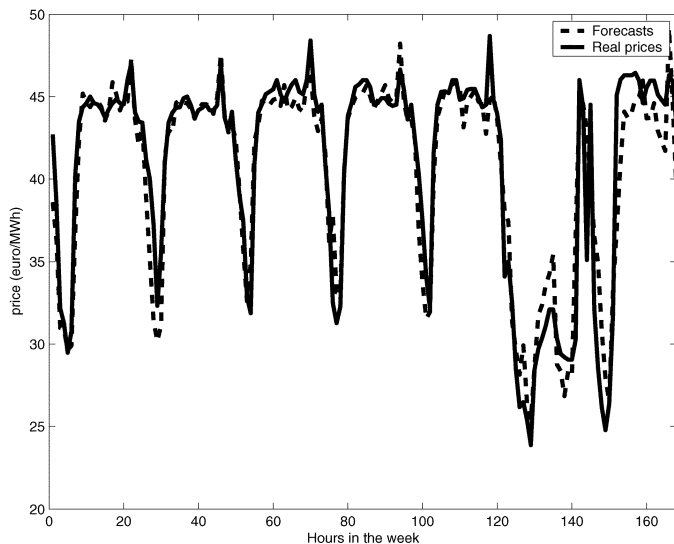


Fig. 13. Real prices and forecasts 15th–21st April 2002.

to the end of 2001 and the beginning of 2002, when the prices were extremely high and volatile, are being taken into account to build the model. Nevertheless, this does not affect negatively the accuracy of the forecasts. Real prices and forecasts are shown in Fig. 13. Prediction errors appear in Table VII.

As we have computed forecasts for every hour in the period under study, Table VIII (which content is illustrated in Fig. 14), includes the quantiles Q_{25} , Q_{50} and Q_{75} of the daily MAPE, depending on the day of the week. Using 80 weeks to build weekday models does not allow the computation of forecasts until the 81st week in the period under study, which means that the first day for which we can compute a forecast is July 21, 1999. However, although the prediction error corresponding to Length equal to 80 weeks is lower, there are not statistically significant differences between using 44 or 80 weeks. Here we provide global results for Length 44 weeks in order to be able to show results for a longer period, since if we use 44 weeks the first week for which we can compute forecasts is the 45th of the period under study (so, 8 November 1998).

The results indicate that the mixed model designed provides accurate forecasts, not only for specific weeks but in general conditions, as this period (1998–2003) includes stages of different levels and variability in prices.

We have developed a “global model” that almost always computes accurate forecasts. Indeed, as the errors we obtain for the six years under study are not higher or even lower than the ones obtained by other authors for specific weeks, it seems that we reached our goal.

VI. CONCLUSIONS

This paper develops an easy methodology for building mixed models for forecasting one-day-ahead electricity prices. Several methods corresponding to different combinations of factor levels (convenience or not of analysing separately the prices in working days and weekends and length of the time series used to forecast) are compared. The mixed model is built combining the advantages of several of these methods.

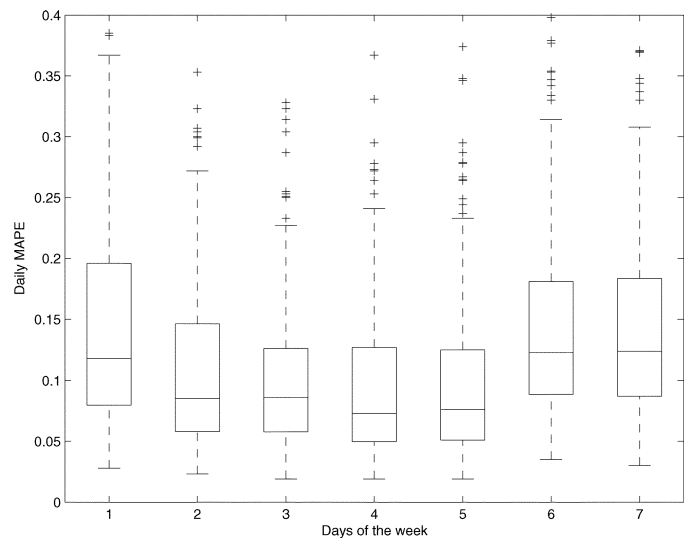


Fig. 14. Boxplot of prediction errors for the whole period considered.

A complete study was carried out to decide between model 24 and 48 and to determine the optimal length of the time series. The analysis, given the size and representativity of the sample (1998–2003), is exhaustive and allows one to draw valid conclusions for price forecasting in the future.

A mixed model to forecast next-day electricity prices is developed. We recommend computing forecasts for working days with model 48 (using only weekdays) and with model 24 for weekends (using complete weeks). We have also determined the optimal length of the time series used to estimate the models for weekends (about 44 weeks) and an appropriate length starting from which, prediction errors do not decrease significantly for weekdays. Splitting the complete time series into 24 hourly time series leads to a much more homogeneous generating processes, which means that adding more information (longer series) allows the computation of better estimations of the parameters of the ARIMA models and this is reflected in terms of prediction errors.

These conclusions have required a great computational effort, as the price for every hour in the period under study has been computed using the 20 models adjusted and estimated for the 20 possible combination of the factors Model (2 levels) and Length (10 levels).

Once this computational effort has been done to design the mixed model we propose, computing a forecast for the 24 hourly prices of tomorrow having the prices of the prior weeks, takes less than 20 seconds. Besides, these forecasts are computed using TRAMO, a free and easy to use software. The model could be useful for different agents interested on having accurate next-day forecasts for electricity prices. Once the mixed model has been proposed no previous knowledge about time series analysis is required for producing a new forecast.

Forecasts have been computed for every hour in the years 1998–2003, and the average error for the whole period is 12.61%. It has been calculated weighting adequately (5:2) the 11.9% obtained for weekdays and 14.4% obtained for weekends. These results reflect the great performance and accuracy

TABLE X
ESTIMATED PARAMETERS

	Estimation of the parameters		
Hour 5	$\theta_1 = -0.47$	$\Theta_1 = -0.99$	
Hour 10	$\theta_1 = -0.37$	$\theta_2 = -0.32$	
Hour 14	$\phi_1 = -0.81$	$\theta_1 = -0.31$	
Hour 18	$\theta_1 = -0.57$		
Hour 22	$\phi_1 = -0.35$	$\theta_1 = -0.80$	

of the new mixed model, not only for a few weeks or days but for a huge and significant sample.

The idea of building mixed models to forecast one-day-ahead electricity prices can also be applied to other electricity markets, like the PJM interconnection, for which Nogales and Conejo [12] obtained by means of ARIMA models and transfer function models forecasting errors of similar magnitude as the ones we have for the Spanish market.

It would be of interest for future research building mixed models for transfer function models. This can be done by taking into account not only the past values of the price in the hour for which we are forecasting, but also past values of the price in the adjacent hours, or including explanatory variables such as demand or temperature using the methodology developed in Cottet and Smith [17] or in Nogales and Conejo [12].

APPENDIX ESTIMATIONS AND DIAGNOSTIC CHECKING

Tables IX and X provide estimations for the 26 May 2000. Results for Hour 5, Hour 10, Hour 14, Hour 18, and Hour 22 are provided. They have been chosen as representative hours, as can be observed in Fig. 2.

As it is a Friday, Model 48 has been used. So the models were built using the previous five-day weeks (considering only weekdays). The notation used was introduced in equation 2.

Table IX also includes the value of the Q-statistic, [18], used to check for remaining correlations between the residuals after fitting the model. The Ljung-Box test is a global one for the h first autocorrelation coefficients.

The Q-statistic for the h first autocorrelation coefficients, $Q(h)$, is asymptotically distributed as a Chi-squared with $h - n$ degrees of freedom, where n is the number of estimated parameters.

In general, we will reject the null hypothesis of uncorrelated residuals when the probability $\Pr((\chi_{h-n}^2) > Q(h))$ is small, usually smaller than 0.05 or 0.01.

ACKNOWLEDGMENT

The authors would like to thank J. Mira for his help and comments, J. Bógalo for his modifications on the interface TS of TRAMO for MatLab, and E. Pinilla for his help with vectorial graphs.

REFERENCES

[1] A. J. Conejo, F. J. Nogales, and J. M. Arroyo, "Price-taker bidding strategy under price uncertainty," *IEEE Trans. Power Syst.*, vol. 17, pp. 1081–1088, 2002.

[2] A. J. Conejo, F. J. Nogales, J. M. Arroyo, and R. García-Bertrand, "Risk constrained self scheduling of a thermal power producer," *IEEE Trans. Power Syst.*, vol. 19, pp. 1569–1574, 2004.

[3] B. Ramsay and A. J. Wang, *A Neural Network Based Estimator for Electricity Spot-Pricing With Particular Reference to Weekend and Public Holidays*, 1998.

[4] B. R. Szkuta, L. A. Sanabria, and T. S. Dillon, "Electricity price short-term forecasting using artificial neural networks," *IEEE Trans. Power Syst.*, vol. 14, 2004.

[5] C. P. Rodríguez and G. J. Anders, "Energy price forecasting in the ontario competitive power system market," *IEEE Trans. Power Syst.*, vol. 19, 2004.

[6] J. D. Nicolaisen, C. W. Richter, and G. B. Sheblé, "Price signal analysis for competitive electric generation companies using artificial neural networks," in *Proc. Conf. Elect. Utility Deregulation and Restructuring and Power Technologies*, 2000.

[7] F. J. Nogales, J. Contreras, A. J. Conejo, and R. Espínola, "Forecasting next-day electricity prices by time series models," *IEEE Trans. Power Syst.*, vol. 17, pp. 342–348, 2002.

[8] J. Contreras, R. Espínola, F. J. Nogales, and A. J. Conejo, "ARIMA models to predict next-day electricity prices," *IEEE Trans. Power Syst.*, vol. 18, pp. 1014–1020, 2003.

[9] A. Troncoso, J. Riquelme, J. Riquelme, A. Gómez, and J. L. Martínez, "A comparison of two techniques for next-day electricity price forecasting," in *Proc. 13th Int. Conf. Database and Expert Systems Applications*, 2002, vol. 2453, Lecture Notes In Computer Science.

[10] J. Crespo-Cuaresma, J. Hlouskova, S. Kossmeier, and M. Obersteiner, "Forecasting electricity spot-prices using linear univariate time-series models," *Appl. Energy*, vol. 77, 2004.

[11] A. J. Conejo, J. Contreras, R. Espínola, and M. A. Plazas, "Forecasting electricity prices for a day-ahead pool-based electric energy market," *Int. J. Forecast.*, 2005.

[12] F. J. Nogales and A. J. Conejo, "Electricity price forecasting through transfer function models," *J. Oper. Res. Soc.*, 2006.

[13] G. Caporello and A. Maravall, "TSW revised reference manual," *Banco de España*, 2004.

[14] W. S. Cleveland, "Robust locally weighted regression and smoothing scatterplots," *J. Amer. Statist. Assoc.*, vol. 74, pp. 829–836, 1979.

[15] D. C. Montgomery, *Design and Analysis of Experiments*. New York: Wiley, 1984.

[16] A. J. Conejo, M. A. Plazas, R. Espínola, and A. B. Molina, "Day-ahead electricity price forecasting using wavelet transform and ARIMA models," *IEEE Trans. Power Syst.*, vol. 20, pp. 1035–1042, 2002.

[17] R. Cottet and M. Smith, "Bayesian modeling and forecasting of intraday electricity load," *J. Amer. Statist. Assoc.*, vol. 98, no. 464, 2003.

[18] G. M. Ljung and G. E. P. Box, "On a measure of lack of fit in time series," *Biometrika*, 1978.

Carolina García-Martos was born in Madrid, Spain, in 1981. She received the degree in industrial engineering from the Universidad Politécnica de Madrid in July 2005. She is now pursuing the Ph.D. degree and is a Teaching Assistant of statistics at Escuela Técnica Superior de Ingenieros Industriales, Universidad Politécnica de Madrid.

Julio Rodríguez is Associate Professor of statistics and econometrics at Universidad Autónoma de Madrid. His areas of interest are time series, multivariate analysis and graphical methods. His research work has appeared in the *Journal of American Statistical Association*, *Journal of Multivariate Analysis*, *Statistica Sinica*, and the *International Journal of Forecasting*, among others.

María Jesús Sánchez received the degree in electrical engineering in 1989 and the Ph.D. degree in applied statistics in 1995, both from the Polytechnical University of Madrid, Spain.

She is Associate Professor of statistics at the Polytechnical University of Madrid. Her current research fields of interest are outliers in time series, kriging models, and reliability of electric power generating systems. research work has appeared in *Reliability Engineering and System Safety* IEEE TRANSACTIONS ON POWER SYSTEMS, *Computational Statistics and I Analysis*, and *Communications in Statistics*, among others.